1250923

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

November 18, 2004

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/505,527*
FILING DATE: *September 23, 2003*
RELATED PCT APPLICATION NUMBER: *PCT/US04/31481*

Certified by

Jon W Dudas

Acting Under Secretary of Commerce
for Intellectual Property
and Acting Director of the U.S.
Patent and Trademark Office

Attorney Docket No.**5470.401PR** PATENT

COVER SHEET FOR FILING PROVISIONAL PATENT APPLICATION (37 CFR §1.51(c)(1))

Mail Stop PROVISIONAL PATENT APPLICATION
Commissioner for Patents
Post Office Box 1450
Alexandria, Virginia 22313-1450

September 23, 2003

This is a request for filing a PROVISIONAL PATENT APPLICATION under 37 C.F.R §1.53(c).

Docket No.	**5470.401PR**
Type a plus sign (+) inside this box ☐	+

INVENTOR(s)/APPLICANT(s)

Name: **Darrel W. Stafford**

Address: Carrboro, North Carolina

TITLE OF THE INVENTION *(280 characters maximum)*
IDENTIFICATION OF THE GENE FOR VITAMIN K EPOXIDE REDUCTASE

ENCLOSED APPLICATION PARTS *(check all that apply)*

☒ Specification (Number of Pages 17)

☒ Drawing(s) (Number of Sheets 2)

☒ Claims (Number of Claims 15)
(A complete provisional application does not require claims 37 C.F.R. §1.51(c)).

☐ Application Data Sheet. See 37 CFR §1.76

☒ Other: Sequence Listing

CORRESPONDENCE ADDRESS

USPTO Customer Number 20792
Myers Bigel Sibley & Sajovec, P.A.
Post Office Box 37428
Raleigh, North Carolina 27627
Telephone: (919) 854-1400
Facsimile: (919) 854-1401

METHOD OF PAYMENT

- ☒ Applicant claims small entity status. See 37 CFR §1.27.
- ☒ Check or money order is enclosed to cover the filing fee.
- ☐ Payment by credit card. Form PTO-2038 is attached.
- ☐ The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account No. 50-0220.

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

- ☐ No.
- ☒ Yes, the name of the U.S. Government agency and the Government contract number are: 5P01 HL06350-42 and 5-R01 HL48318 from the National Institutes of Health

Respectfully submitted,

Kenneth D. Sibley
Registration No. 31,665

CERTIFICATE OF EXPRESS MAIL

Express Mail Label Number EV318420117US
Date of Deposit: September 23, 2003

I hereby certify that this correspondence is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to Mail Stop PROVISIONAL PATENT APPLICATION, Commissioner for Patents, Post Office Box 1450, Alexandria, Virginia 22313-1450.

Vickie Diane Prior

IDENTIFICATION OF THE GENE FOR VITAMIN K EPOXIDE REDUCTASE

Government Support

The present invention was made, in part, with the support of grant numbers 5P01 HL06350-42 and 5-R01 HL48318 from the National Institutes of Health. The United States government has certain rights to this invention.

Field of the Invention

The present invention concerns isolated nucleic acids, host cells containing the same, and methods of use thereof.

Background of the Invention

The function of numerous proteins requires the modification of multiple glutamic acid residues to γ-carboxyglutamate. Among these vitamin K-dependent (VKD) coagulation proteins, FIX (Christmas factor), FVII, and prothrombin are the best known. The observation that a knock-out of the gene for matrix Gla protein results in calcification of the mouse's arteries (Luo, G. et al. (1997) Spontaneous calcification of arteries and cartilage in mice lacking matrix GLA protein. *Nature* 386, 78-81) emphasizes the importance of the vitamin K cycle for proteins with functions other than coagulation. Moreover, Gas6 [5] and other Gla proteins of unknown function[6,7] are expressed in neural tissue and warfarin exposure *in utero* results in mental retardation and facial abnormalities[8-10]. This is consistent with the observation that the expression of VKD carboxylase, the enzyme that accomplishes the Gla modification[11], is temporally regulated in a tissue-specific manner with high expression in the nervous system during early embryonic stages. Concomitant with carboxylation, reduced vitamin K, a co-substrate of the reaction, is converted to vitamin K epoxide. Because the amount of vitamin K in the human diet is limited, vitamin K epoxide (VKOR) must be converted back to vitamin K by VKOR to prevent its depletion. Warfarin, the most widely used anticoagulation drug, targets VKOR and prevents the regeneration of vitamin K. The consequence is a decrease in the concentration of reduced vitamin K, which

results in a reduced rate of carboxylation by the γ-glutamyl carboxylase and in the production of undercarboxylated vitamin K-dependent proteins.

VKOR activity was first reported by Zimmerman and Matschiner in 1974[12]. Since this first report, several groups, including ours, have attempted its purification with little success.

Summary of the Invention

A first aspect of the present invention is an isolated nucleic acid encoding vitamin K epoxide reductase (VKOR), particularly mammalian (*e.g.*, human, ovine, bovine, monkey, etc.) VKOR. Examples include *(a)* nucleic acids as disclosed herein, such as isolated nucleic acids having the sequence given in SEQ ID NO: 8 or SEQ ID NO: 9; *(b)* nucleic acids that hybridize to isolated nucleic acids of (a) above or the complement thereof (e.g., under stringent conditions), and/or have substantial sequence identity to nucleic acids of (a) above (*e.g.*, are 80, 85, 90 95 or 99% identical to nucleic acids of (a) above), and encode a VKOR; and *(c)* nucleic acids that differ from the nucleic acids of (a) or (b) above due to the degeneracy of the genetic code, but code for a VKOR encoded by a nucleic acid of (a) or (b) above.

A second aspect of the present invention is a recombinant nucleic acid comprising a nucleic acid encoding vitamin K epoxide reductase as described herein operatively associated with a heterologous promoter.

A further aspect of the present invention is a host cell that contains and expresses a recombinant nucleic acid as described above. Suitable host cells include plant, animal, mammal, insect, yeast and bacterial cells.

A further aspect of the present invention is an oligonucleotide that hybridizes to an isolated nucleic acid encoding VKOR as described herein.

A further aspect of the present invention is isolated and purified VKOR (*e.g.*, VKOR purified to homogeneity) encoded by a nucleic acid as described herein.

A further aspect of the present invention is, a method of making a vitamin K dependent protein which comprises culturing a host cell which expresses a vitamin K dependent protein in the presence of vitamin K, and then harvesting the vitamin K dependent protein from the culture, the host cell containing and expressing a heterologous nucleic acid encoding vitamin K

dependent carboxylase, characterized by the host cell further containing and expressing a heterologous nucleic acid encoding vitamin K epoxide reductase (VKOR) as described herein. Preferably the expression of the VKOR causes the cell to express greater levels of the vitamin K dependent protein than would be expressed in the absence of expression of the VKOR. Any desired vitamin K dependent protein may be expressed, including but not limited to Factor VII, Factor IX, Factor X, Protein C, Protein S, and Prothrombin. Any host cell may be used as described above, although in some embodiments non-human or even non-mammalian host cells are preferred.

Brief Description of the Drawings

Figure 1. For each of the 13 siRNA pools, three T7 flasks containing A549 cells were transfected and VKOR activity determined after 72 h. The VKOR assay used 25 μM vitamin K epoxide. One siRNA pool specific for gene gi:13124769 reduced VKOR activity by 64%-70% in eight repetitions.

Figure 2. Time course of inhibition of VKOR activity by the siRNA pool specific for gi:13124769 in A549 cells. VKOR activity decreased continuously during this time period while the level of its mRNA decreased rapidly to about 20% of normal. 25 μM vitamin K epoxide was used for this assay. The siRNA does not affect the activity of VKD carboxylase or the level of lamin A/C mRNA.

Figure 3. VKOR activity was detected when mGC_11276 was expressed in Sf9 insect cells. ~$1x10^6$ cells were used in this assay. Reactions were performed using 32 μM KO at 30°C for 30 minutes in Buffer D. Blank Sf9 cells served as a negative control and A549 cells as a reference.

Figure 4. Inhibition of VKOR by warfarin. Reactions were performed using 1.6 mg microsomal proteins made from VKOR_Sf9 cells, 60 μM KO, and various concentration of warfarin at 30°C for 15 minutes in Buffer D.

Detailed Description of the Preferred Embodiments

The present invention is explained in greater detail below. This description is not intended to be a detailed catalog of all the different ways in which the invention may be implemented, or all the features that may be added to the instant invention. For example, features illustrated with respect to one embodiment may be incorporated into other embodiments, and features illustrated with respect to a particular embodiment may be deleted from that embodiment. In addition, numerous variations and additions to the various embodiments suggested herein will be apparent to those skilled in the art in light of the instant disclosure which do not depart from the instant invention. Hence, the following specification is intended to illustrate some particular embodiments of the invention, and not to exhaustively specify all permutations, combinations and variations thereof.

The "Sequence Listing" attached hereto forms a part of the instant specification as if fully set forth herein.

The present invention may be carried out based on the instant disclosure and further utilizing methods, components and features known in the art, including but not limited to those described in US Patent No. 5,268,275 to Stafford and Wu and US Patent No. 6,531,298 to Stafford and Chang, the disclosures of which are to be incorporated by reference herein in their entirety as if fully set forth.

As used herein, "nucleic acids" encompass both RNA and DNA, including cDNA, genomic DNA, and synthetic (e.g., chemically synthesized) DNA. The nucleic acid may be double-stranded or single-stranded. Where single-stranded, the nucleic acid may be a sense strand or an antisense strand. The nucleic acid may be synthesized using oligonucleotide analogs or derivatives (e.g., inosine or phosphorothioate nucleotides). Such oligonucleotides can be used, for example, to prepare nucleic acids that have altered base-pairing abilities or increased resistance to nucleases.

An "isolated nucleic acid" is a DNA or RNA that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (one on the 5' end and one on the 3' end) in the naturally occurring genome of the organism from which it is derived. Thus, in one embodiment,

an isolated nucleic acid includes some or all of the 5' non-coding (e.g., promoter) sequences that are immediately contiguous to the coding sequence. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA that is part of a hybrid gene encoding an additional polypeptide sequence. The term "isolated" can refer to a nucleic acid or polypeptide that is substantially free of cellular material, viral material, or culture medium (when produced by recombinant DNA techniques), or chemical precursors or other chemicals (when chemically synthesized). Moreover, an "isolated nucleic acid fragment" is a nucleic acid fragment that is not naturally occurring as a fragment and would not be found in the natural state.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20 to 25 nucleotides, which can be used in PCR amplification or a hybridization assay, or a microarray. Oligonucleotides may be natural or synthetic, e.g., DNA, RNA, modified backbones, etc.

Where a particular nucleotide sequence is said to have a specific percent identity to a reference nucleotide sequence of a defined length, the percent identity is relative to the reference nucleotide sequence. Thus, a nucleotide sequence that is 50%, 75%, 85%, 90%, 95% or 99% identical to a reference nucleotide sequence that is 100 bases long can be a 50, 75, 85, 90, 95 or 99 bases that are completely identical to a 50, 75, 85, 90, 95 or 99 nucleotide sequence of the reference nucleotide sequence. It also might be a 100 bases in length nucleotide sequence which is 50%, 75%, 85%, 90%, 95% or 99% identical to the reference nucleotide sequence over its entire length. Of course, there are other nucleotide sequences that will also meet the same criteria.

A nucleic acid sequence that is "substantially identical" to a VKOR nucleotide sequence is at least 80%, 85% 90%, 95% or 99% identical to the nucleotide sequence of **SEQ ID NO: 9**. For purposes of comparison of nucleic acids, the length of the reference nucleic acid sequence will generally be at least 40 nucleotides, e.g., at least 60 nucleotides or more nucleotides. Sequence identity can be measured using sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705).

The VKOR polypeptides of the invention include, but are not limited to, recombinant polypeptides and natural polypeptides. The invention also encompasses nucleic acid sequences that encode forms of VKOR polypeptides in which naturally occurring amino acid sequences are altered or deleted. Preferred nucleic acids encode polypeptides that are soluble under normal physiological conditions. Also within the invention are nucleic acids encoding fusion proteins in which a portion of VKOR is fused to an unrelated polypeptide (e.g., a marker polypeptide or a fusion partner) to create a fusion protein. For example, the polypeptide can be fused to a hexa-histidine tag to facilitate purification of bacterially expressed polypeptides, or to a hemagglutinin tag to facilitate purification of polypeptides expressed in eukaryotic cells, or to an HPC4 tag to facilitate purification of polypeptides by affinity chromatography or immunoprecipitation. The invention also includes isolated, for example, polypeptides (and the nucleic acids that encode these polypeptides) that include a first portion and a second portion; the first portion includes, e.g., a VKOR polypeptide, and the second portion includes, e.g., a detectable marker.

The fusion partner can be, for example, a polypeptide which facilitates secretion, e.g., a secretory sequence. Such a fused polypeptide is typically referred to as a preprotein. The secretory sequence can be cleaved by the host cell to form the mature protein. Also within the invention are nucleic acids that encode VKOR fused to a polypeptide sequence to produce an inactive

preprotein. Preproteins can be converted into the active form of the protein by removal of the inactivating sequence.

The invention also includes nucleic acids that hybridize, e.g., under stringent hybridization conditions (as defined herein) to all or a portion of the nucleotide sequence of **SEQ ID NOS: 1–6, 8** or **9** or their complements. The hybridizing portion of the hybridizing nucleic acids is typically at least 15 (e.g., 20, 30, or 50) nucleotides in length. The hybridizing portion of the hybridizing nucleic acid is at least 80%, e.g., at least 95%, or at least 98%, identical to the sequence of a portion or all of a nucleic acid encoding a VKOR polypeptide. Hybridizing nucleic acids of the type described herein can be used as a cloning probe, a primer (e.g., a PCR primer), or a diagnostic probe. Also included within the invention are small inhibitory RNAs (siRNAs) that inhibit the function of VKOR, as determined, for example, in an activity assay.

In another embodiment, the invention features cells, e.g., transformed host cells, that contain a nucleic acid is encompassed by the invention. A "transformed cell" is a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a nucleic acid encoding an VKOR polypeptide. Both prokaryotic and eukaryotic cells are included, e.g., bacteria, yeast, insect, mouse, rat, human and the like.

The invention also features genetic constructs (e.g., vectors and plasmids) that include a nucleic acid of the invention which is operably linked to a transcription and/or translation sequence to enable expression, e.g., expression vectors. By "operably linked" is meant that a selected nucleic acid, e.g., a DNA molecule encoding a VKOR polypeptide, is positioned adjacent to one or more sequence elements, e.g., a promoter, which directs transcription and/or translation of the sequence such that the sequence elements can control transcription and/or translation of the selected nucleic acid.

The invention also features purified or isolated VKOR polypeptides. As used herein, both "protein" and "polypeptide" mean any chain of amino acids, regardless of length or post-translational modification (e.g., glycosylation,

phosphorylation or N-myristoylation). Thus, the terms "VKOR polypeptide" include full-length, naturally occurring VKOR proteins, respectively, as well as recombinantly or synthetically produced polypeptides that correspond to a full-length, naturally occurring VKOR protein, or to a portion of a naturally occurring or synthetic VKOR polypeptide.

A "purified" or "isolated" compound or polypeptide is a composition that is at least 60% by weight the compound of interest, e.g., a VKOR polypeptide or antibody that is separated or substantially free from at least some of the other components of the naturally occurring organism or virus, for example, the cell or viral structural components or other polypeptides or nucleic acids commonly found associated with the polypeptide. As used herein, the "isolated" polypeptide is at least about 25%, 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% or more pure (w/w). Preferably the preparation is at least 75% (e.g., at least 90% or 99%) by weight the compound of interest. Purity can be measured by any appropriate standard method, e.g., column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

Preferred VKOR polypeptides include a sequence substantially identical to all or a portion of a naturally occurring VKOR polypeptide. Polypeptides "substantially identical" to the VKOR polypeptide sequences described herein have an amino acid sequence that is at least 80% or 85% (e.g., 90%, 95% or 99%) identical to the amino acid sequence of the VKOR polypeptides of **SEQ ID NO: 10**. For purposes of comparison, the length of the reference VKOR polypeptide sequence will generally be at least 16 amino acids, e.g., at least 20 or 25 amino acids.

In the case of polypeptide sequences that are less than 100% identical to a reference sequence, the non-identical positions are preferably, but not necessarily, conservative substitutions for the reference sequence. Conservative substitutions typically include but are not limited to substitutions within the following groups: glycine and alanine; valine, isoleucine, and

leucine; aspartic acid and glutamic acid; asparagine and glutamine; serine and threonine; lysine and arginine; and phenylalanine and tyrosine.

Where a particular polypeptide is said to have a specific percent identity to a reference polypeptide of a defined length, the percent identity is relative to the reference polypeptide. Thus, a polypeptide that is 50%, 75%, 85%, 90%, 95% or 99% identical to a reference polypeptide that is 100 amino acids long can be a 50, 75, 85, 90, 95 or 99 amino acid polypeptide that is completely identical to a 50, 75, 85, 90, 95 or 99 amino acid long portion of the reference polypeptide. It also might be a 100 amino acid long polypeptide which is 50%, 75%, 85%, 90%, 95% or 99% identical to the reference polypeptide over its entire length. Of course, other polypeptides also will meet the same criteria.

The invention also features purified or isolated antibodies that specifically bind to a VKOR polypeptide. By "specifically binds" is meant that an antibody recognizes and binds a particular antigen, e.g., a VKOR polypeptide, but does not substantially recognize and bind other molecules in a sample. In one embodiment the antibody is a monoclonal antibody.

In another aspect, the invention features a method for detecting a VKOR polypeptide in a sample. This method includes: obtaining a sample suspected of containing VKOR; contacting the sample with an antibody that specifically binds a VKOR polypeptide under conditions that allow the formation of complexes of an antibody and VKOR; and detecting the complexes, if any, as an indication of the presence of VKOR in the sample.

Also encompassed by the invention is a method of obtaining a gene related to (i.e., a functional homologue of) the VKOR gene. Such a method entails obtaining a labeled probe that includes an isolated nucleic acid which encodes all or a portion of VKOR, or a homolog thereof; screening a nucleic acid fragment library with the labeled probe under conditions that allow hybridization of the probe to nucleic acid fragments in the library, thereby forming nucleic acid duplexes; isolating labeled duplexes, if any; and

preparing a full-length gene sequence from the nucleic acid fragments in any labeled duplex to obtain a gene related to the VKOR gene.

The present invention and the various methods and compounds therein are explained in greater detail in the following non-limiting examples.

EXAMPLE 1

siRNA design and synthesis

siRNAs were selected using an advanced version of a rational design algorithm (Reynolds, A. (2003) RNA interference: mechanistic implications and rational siRNA design. *Nature Biotechnology* Submitted). For each of the 13 genes, four siRNAs duplexes with the highest scores were selected and a BLAST search was conducted using the Human EST database. To minimize the potential for off-target silencing effects, only those sequence targets with more than three mismatches against un-related sequences were selected (Jackson, A. L. et al. (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* 21, 635-7). All duplexes were synthesized in Dharmacon (Lafayette, CO) as 21-mers with UU overhangs using a modified method of 2'-ACE chemistry (Scaringe, S. A. (2000) Advanced 5'-silyl-2'-orthoester approach to RNA oligonucleotide synthesis. *Methods Enzymol* 317, 3-18) and the AS strand was chemically phosphorylated to insure maximum activity (Martinez, J. et al. (2002) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* 110, 563-74).

EXAMPLE 2

siRNA transfection

Transfection was essentially as described by Harborth et al. (Harborth, J. et al. (2001) Identification of essential genes in cultured mammalian cells using small interfering RNAs. *J Cell Sci* 114, 4557-65) with minor modifications.

EXAMPLE 3

VKOR activity assay

siRNA transfected A549 cells were trypsinized and washed twice with cold PBS. $1.5x10^7$ cells were taken for each VKOR assay. 200 µL buffer D (250 mM Na_2HPO_4-NaH_2PO_4, 500 mM KCl, 20% glycerol and 0.75% CHAPS, pH 7.4) was added to the cell pellet followed by sonication of the cell lysate. For assays of solubilized microsomes, microsomes were prepared from $2x10^9$ cells as described (Lin, P. J. et al. (2002) The putative vitamin K-dependent gamma-glutamyl carboxylase internal propeptide appears to be the propeptide binding site. *J Biol Chem* 277, 28584-91); 10 to 50 µL of solubilized microsomes were used for each assay. Vitamin K epoxide was added to the concentration indicated in the figure legends and DTT was added to 4 mM to initiate the reaction. The reaction mixture was incubated in yellow light at 30°C for 30 minutes and stopped by adding 500 µL 0.05 M $AgNO_3$: isopropanol (5:9). 500 µL hexane was added and the mixture was vortexed vigorously for 1 minute to extract the vitamin K and KO. After 5 minutes centrifugation the upper organic layer was transferred to a 5-mL brown vial and dried with N_2. 150 µL HPLC buffer, acetonitrile:isopropanol:water (100:7:2), was added to dissolve the vitamin K and KO and the sample was analyzed by HPLC on an A C-18 column (Vydac, cat#218TP54).

EXAMPLE 4

RT-qPCR (reverse transcriptase quantitative PCR)

$1x10^6$ cells were washed with PBS twice and total RNA was isolated with Trizol reagent according to the manufacturers protocol (Invitrogen). 1µg of RNA was digested by RQ1 DNaseI (Promega) and heat-inactivated. First strand cDNA was made with M-MLV reverse transcriptase (Invitrogen). cDNAs were mixed with DyNAmo SYBR Green qPCR pre-mix (Finnzymes) and real-time PCR was performed with an Opticon II PCR thermal cycler (MJ Research). The following primers were used:

13124769-5' (F): (TCCAACAGCATATTCGGTTGC, **SEQ ID NO: 1**);
13124769-3 (R)': (TTCTTGGACCTTCCGGAAACT, **SEQ ID NO: 2**);
GAPDH-F: (GAAGGTGAAGGTCGGAGTC, **SEQ ID NO: 3**);
GAPDH-R: (GAAGATGGTGATGGGATTTC, **SEQ ID NO: 4**);
Lamin-RT-F: (CTAGGTGAGGCCAAGAAGCAA, **SEQ ID NO: 5**) and

Lamin-RT-R: (CTGTTCCTCTCAGCAGACTGC, **SEQ ID NO: 6**).

EXAMPLE 5

Over-expression of VKOR in Sf9 insect cell line

The cDNA for mGC11276 coding region was cloned into pVL1392 (Pharmingen), with the HPC4 tag (EDQVDPRLIDGK, **SEQ ID NO: 7**) at its amino terminus and expressed in Sf9 cells as described (Li, T. et al. (2000) Identification of a Drosophila vitamin K-dependent gamma-glutamyl carboxylase. *J Biol Chem* 275, 18291-6).

EXAMPLE 6

Gene selection

Two recent reports suggesting the chromosomal location of the gene for VKOR stimulated us to attempt a different approach to identify the gene for this important enzyme. First, Kohn and Pelz (Kohn, M. H. & Pelz, H. J. A. (2000) Gene-anchored map position of the rat warfarin-resistance locus, Rw, and its orthologs in mice and humans. *Blood* 96, 1996-8) mapped warfarin resistance (*Rw*) to rat chromosome one between markers D1Mit13 and D1Rat67. In this region four marker genes, Myl2, Itgam, IL4Ra and Fgfr2, allowed Kohn and Pelz to locate the syntenic regions of mouse and human chromosomes. They reported that these four markers are found on mouse chromosome seven 60cM-63cM, and on human chromosomes ten, twelve, and sixteen. Secondly, Fregin *et al.* (Fregin, A. et al. (2002) Homozygosity mapping of a second gene locus for hereditary combined deficiency of vitamin K-dependent clotting factors to the centromeric region of chromosome 16. *Blood* 100, 3229-32) mapped a gene for combined deficiencies of VKD proteins to human chromosome16p12-q21. We found three, Myl2, Itgam and IL4Ra, of the four marker genes on human chromosome 16p11-13. Because warfarin resistance and combined deficiency of VKD proteins map to the same region, we focused our search on human chromosome sixteen between markers D16S3131 and D16S419. This region corresponds to chromosome 16 at 50cM-65cM on the genetic map and 26-46.3Mb on the physical map. We analyzed the 190 predicted coding sequences in this region by a BLASTX search of the NCBI non-redundant protein database. Those human genes and

orthologs from related species with known function were eliminated. Because VKOR appears to be a transmembrane protein (Carlisle, T. L. & Suttie, J. W. (1980) Vitamin K dependent carboxylase: subcellular location of the carboxylase and enzymes involved in vitamin K metabolism in rat liver. *Biochemistry* 19, 1161-7), the remaining genes were translated according to the cDNA sequences in the NCBI database and analyzed with the programs TMHMM and TMAP (Biology WorkBench, San Diego Supercomputer System) to predict those with transmembrane domains. Thirteen genes predicted to code for integral membrane proteins were chosen for further analysis.

EXAMPLE 7

Cell line screening for VKOR activity

Our strategy was to identify a cell line expressing relatively high amounts of VKOR activity and use siRNA to systematically knockdown all thirteen candidate genes. siRNA, double stranded RNA of 21-23 nucleotides, has been shown to cause specific RNA degradation in cell culture (Hara, K. et al. (2002) Raptor, a binding partner of target of rapamycin (TOR), mediates TOR action. *Cell* 110, 177-89, Krichevsky, A. M. & Kosik, K. S. (2002) RNAi functions in cultured mammalian neurons. *Proc Natl Acad Sci U S A* 99, 11926-9, Burns, T. F. et al. (2003) Silencing of the Novel p53 Target Gene Snk/Plk2 Leads to Mitotic Catastrophe in Paclitaxel (Taxol)-Exposed Cells. *Mol Cell Biol* 23, 5556-71). However, application of siRNA for large scale screening in mammalian cells has not previously been reported because of the difficulty in identifying a functional target for a specific mammalian cell mRNA (Holen, T. et al. (2003) Similar behaviour of single-strand and double-strand siRNAs suggests they act through a common RNAi pathway. *Nucleic Acids Res* 31, 2401-7). The recent development of a rational selection algorithm (Reynolds, A. (2003) RNA interference: mechanistic implications and rational siRNA design. *Nature Biotechnology* Submitted) for siRNA design increases the probability that a specific siRNA can be developed; furthermore, the probability of success can be increased by pooling four rationally selected siRNAs. Using siRNA to search for previously unidentified genes has the advantage that, even if VKOR activity requires the product of more than one

gene for activity, the screen should still be effective because the assay determines the loss of enzymatic activity.

We screened fifteen cell lines and found a human lung carcinoma line, A549, which exhibited sufficient warfarin-sensitive VKOR activity for facile measurement. A second human colorectal adenocarcinoma cell line, HT29, that expressed very little VKOR activity was used as a reference.

EXAMPLE 8

siRNA inhibition of VKOR activity in A549 cells

We transfected each of the thirteen pools of siRNA in triplicate into A549 cells and after 72 hours assayed for VKOR activity. One siRNA pool specific for gene gi:13124769 reduced VKOR activity by 64%-70% in eight separate assays (Fig. 1).

One possible reason that VKOR activity was inhibited to only ~35% of its initial activity after 72 hours is that the half-life of mammalian proteins varies greatly (from minutes to days) (Zhang, W. et al. (1996) The major calpain isozymes are long-lived proteins. Design of an antisense strategy for calpain depletion in cultured cells. *J Biol Chem* 271, 18825-30, Bohley, P. (1996) Surface hydrophobicity and intracellular degradation of proteins. *Biol Chem 377*, 425-35, Dice, J. F. & Goldberg, A. L. (1975) Relationship between in vivo degradative rates and isoelectric points of proteins. *Proc Natl Acad Sci U S A* 72, 3893-7), and we are inhibiting mRNA translation, not enzyme activity. Therefore, we carried the cells through eleven days and followed their VKOR activity. Figure 2 shows that the level of mRNA for gi:13124769 mRNA decreased rapidly to about 20% of normal while VKOR activity decreased continuously during this time period. This reduction in activity is not a general effect of the siRNA or the result of cell death because the level of VKD carboxylase activity and lamin A/C mRNA remained constant. Furthermore, the level of gi:132124769 mRNA is four fold lower in HT-29 cells, which have low VKOR ativity, than in A549 cells that exhibit high VKOR activity (data not shown. These data indicate that gi:13124769 corresponds to the VKOR gene.

EXAMPLE 9

Identification of gene encoding VKOR

The gene, IMAGE 3455200 (gi:13124769, **SEQ ID NO: 8**), identified by us to encode VKOR, maps to human chromosome 16p11.2, mouse chromosome 7F3, and rat chromosome 1:180.8 Mb. There are 338 cDNA clones in the NCBI database representing seven different splicing patterns (NCBI AceView program). These are composed of all or part of two to four exons. Among these, the most prevalent isoform, mGC11276, has three exons and is expressed at high levels in lung and liver cells. This three exon transcript (**SEQ ID NO: 9**) encodes a predicted protein of 163 amino acids with a mass of 18.2 kDa (**SEQ ID NO: 10**). It is a putative N-myristoylated endoplasmic reticulum protein with one to three transmembrane domains, depending upon the program used for prediction. It has seven cysteine residues, which is consistent with observations that the enzymatic activity is dependent upon thiol reagents (Thijssen, H. H. et al. (1994) Microsomal lipoamide reductase provides vitamin K epoxide reductase with reducing equivalents. *Biochem J* 297 (Pt 2), 277-80). Five of the seven cysteines are conserved among human, mice, rat, zebrafish, *Xenopus* and *Anopheles*.

To confirm that we had identified the VKOR gene, we expressed the most prevalent form of the enzyme (three exons) in *Spodoptera frugiperda*, Sf9 cells. Sf9 cells have no measurable VKOR activity but exhibit warfarin sensitive activity when transfected with mGC11276 cDNA (Figure 3). VKOR activity is observed from constructs with an epitope tag at either their amino or carboxyl terminus. This tag should assist in the purification of VKOR.

VKOR should exhibit warfarin sensitivity, therefore we made microsomes from Sf9 cells expressing VKOR and tested for warfarin sensitivity. The VKOR activity is warfarin-sensitive (Figure 4).

In summary, we report here the first example of using siRNA in mammalian cells to identify an unknown gene. The identity of the VKOR gene was confirmed by its expression in insect cells. The VKOR gene encodes several isoforms. It will be important to characterize the activity and expression pattern of each isoform. Millions of people world-wide utilize warfarin to inhibit coagulation; therefore it is important to further characterize VKOR as it may lead to more accurate dosing or design of safer, more effective, anti-coagulants.

The foregoing is illustrative of the present invention, and is not to be construed as limiting thereof. The invention is defined by the following claims, with equivalents of the claims to be included therein.

THAT WHICH IS CLAIMED IS:

1. An isolated nucleic acid encoding vitamin K epoxide reductase (VKOR).

2. The isolated nucleic acid according to claim 1 encoding mammalian VKOR.

3. The isolated nucleic acid according to claim 1 encoding human VKOR.

4. The isolated nucleic acid of claim 1, wherein said nucleic acid is a cDNA.

5. The isolated nucleic acid of claim 1, wherein said nucleic acid is a genomic DNA.

6. A recombinant nucleic acid comprising a nucleic acid encoding vitamin K epoxide reductase operatively associated with a heterologous promoter.

7. A host cell that contains and expresses a recombinant nucleic acid according to claim 6.

8. The host cell of claim 7, wherein said host cell is a plant, insect, bacterial or yeast cell.

9. An oligonucleotide that hybridizes to an isolated nucleic acid according to claim 1.

10. The oligonucleotide of claim 9, wherein said oligonucleotide is from 6 to 60 nucleotides in length.

11. In a method of making a vitamin K dependent protein which comprises culturing a host cell which expresses a vitamin K dependent protein in the presence of vitamin K, and then harvesting said vitamin K dependent protein from the culture, said host cell containing and expressing a heterologous nucleic acid encoding vitamin K dependent carboxylase, the improvement comprising:

employing as said host cell a host cell further containing and expressing a heterologous nucleic acid encoding vitamin K epoxide reductase (VKOR).

12. The method according to claim 11, wherein said vitamin K dependent protein is selected from the group consisting of Factor VII, Factor IX, Factor X, Protein C, Protein S, and Prothrombin.

13. The method according to claim 11, wherein said host cell is a plant cell.

14. The method according to claim 11, wherein said host cell is an insect cell.

15. The method according to claim 11, wherein said vitamin K-dependent carboxylase is bovine vitamin K dependent carboxylase.

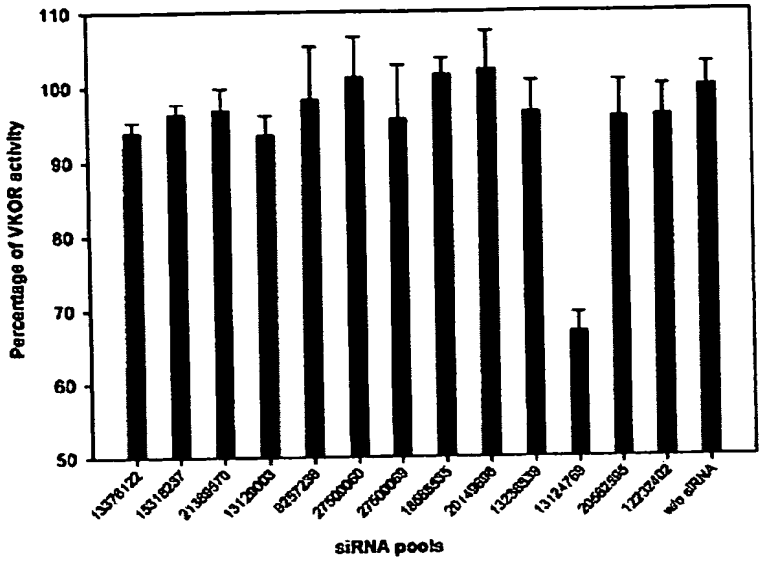

Figure 1

5

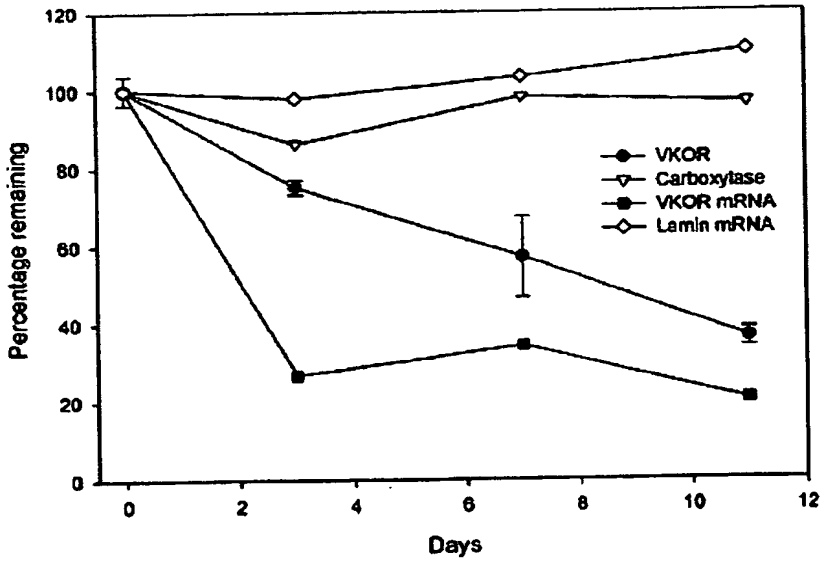

Figure 2

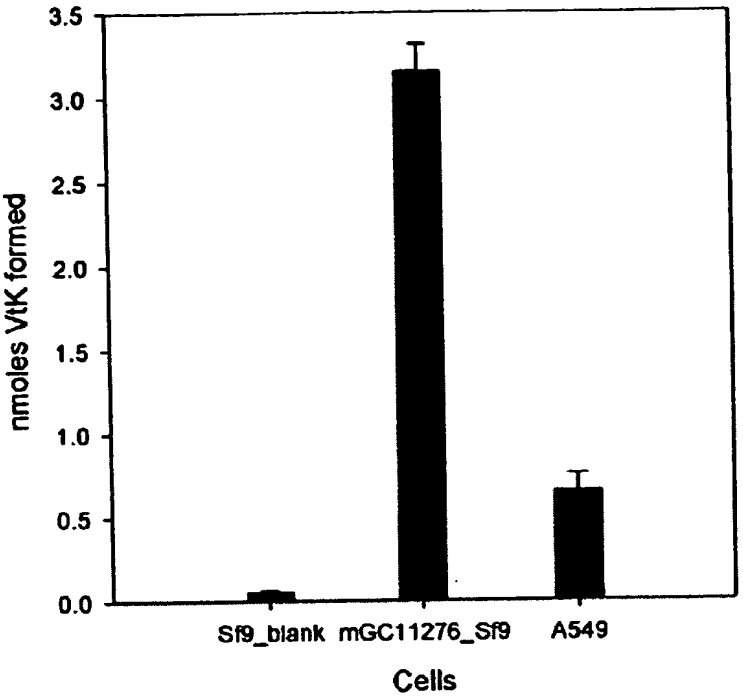

Figure 3

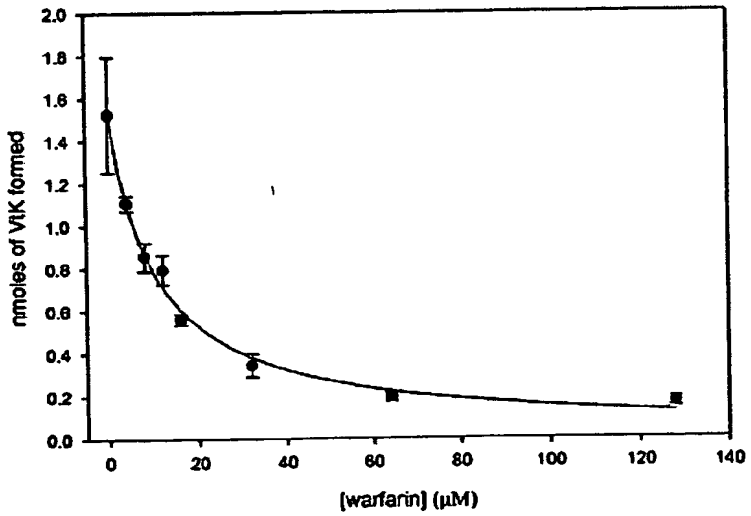

Figure 4

SEQUENCE LISTING

<110> Stafford, Darrel

<120> IDENTIFICATION OF THE GENE FOR VITAMIN K EPOXIDE REDUCTASE

<130> 5470.401

<160> 10

<170> PatentIn version 3.2

<210> 1
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide primer

<400> 1
tccaacagca tattcggttg c 21

<210> 2
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide primer

<400> 2
ttcttggacc ttccggaaac t 21

<210> 3
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotidde primer

<400> 3
gaaggtgaag gtcggagtc 19

<210> 4
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide primer

<400> 4
gaagatggtg atgggatttc 20

<210> 5

<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide primer

<400> 5
ctaggtgagg ccaagaagca a 21

<210> 6
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic oligonucleotide primer

<400> 6
ctgttcctct cagcagactg c 21

<210> 7
<211> 12
<212> PRT
<213> Artificial sequence

<220>
<223> HPC4 tag sequence

<400> 7

Glu Asp Gln Val Asp Pro Arg Leu Ile Asp Gly Lys
1 5 10

<210> 8
<211> 3915
<212> DNA
<213> Homo sapiens

<400> 8
ggttttctcc gcgggcgcct cgggcggaac ctggagataa tgggcagcac ctgggggagc 60

cctggctggg tgcggctcgc tctttgcctg acgggcttag tgctctcgct ctacgcgctg 120

cacgtgaagg cggcgcgcgc ccgggaccgg gattaccgcg cgctctgcga cgtgggcacc 180

gccatcagct gttcgcgcgt cttctcctcc aggtgtgcac gggagtggga ggcgtggggc 240

ctcggagcag ggcggccagg atgccagatg attattctgg agtctgggat cggtgtgccc 300

ggggaacgga cacggggctg gactgctcgc ggggtcgttg cacaggggct gagctaccca 360

gcgatactgg tgttcgaaat aagagtgcga ggcaagggac cagacagtgc tggggactgg 420

gattattccg gggactcgca cgtgaattgg atgccaagga ataacggtga ccaggaaagg 480

cggggaggca ggatggcggt agagattgac gatggtctca aggacggcgc gcaggtgaag 540

ggggggtgttg gcgatggctg cgcccaggaa caaggtggcc cggtctggct gtgcgtgatg 600

gccaggcgtt agcataatga cggaatacag aggaggcgag tgagtggcca gggagctgga 660

gattctgggg tccagggcaa agataatctg cccccgactc ccagtctctg atgcaaaacc 720

gagtgaaccg ttataccagc cttgccattt taagaattac ttaagggccg ggcgcggtgg 780

cccactcctg taatcccagc actttgggag gccgaggcgg atggatcact tgaagtcagg 840

agttgaccag cctggccaac atggtgaaag cctgtctcta ccaaaaatag aaaaattaat 900

cgggcgctat ggcgggtgcc ttaatcccag ctactcgggg gggctaaggc aggagaatcg 960

cttgaacccg ggaggcggag gtttcagtga gccgagatcg cgccactgca ctccagcctg 1020

ggccagagtg agactccgtc tcaaaaaaaa aaaaaaaaaa aaaaaaaaag agacttactt 1080

aaggtctaag atgaaaagca gggcctacgg agtagccacg tccgggcctg gtctggggag 1140

aggggaggat agggtcagtg acatggaatc ctgacgtggc caaaggtgcc cggtgccagg 1200

agatcatcga cccttggact aggatgggag gtcggggaac agaggatagc ccaggtggct 1260

tcttggaaat caccttttctc gggcagggtc caaggcactg ggttgacagt cctaacctgg 1320

ttccacccca ccccacccct ctgccaggtg gggcaggggt ttcgggctgg tggagcatgt 1380

gctgggacag gacagcatcc tcaatcaatc caacagcata ttcggttgca tcttctacac 1440

actacagcta ttgttaggtg agtggctccg ccccctcccc gcccgccccg ccccgccccc 1500

catccccctt ggtcagctca gccccactcc atgcaatctt ggtgatccac acagctgaca 1560

gccagctagc tgctcatcac ggagcgtcct gcgggtgggg atgtggggag gtaactaaca 1620

ggagtctttt aattggttta agtactgtta gaggctgaag ggcccttaaa gacatcctag 1680

gtccccaggt tttttgtttg ttgttgtttt gagacagggt ctggctctgt tgcccaaagt 1740

gaggtctagg atgcccttag tgtgcactgg cgtgatctca gttcatggca acctctgcct 1800

ccctgcccaa gggatcctcc caccttagcc tcccaagcag ctggaatcac aggcgtgcac 1860

cactatgccc agctaatttt tgtttttgtt ttttttggt agagatggtg tctcgccatg 1920

ttgcccaggc tggtctcaag caatctgtct gcctcagcct cccaaagtgc tggggggatt 1980

acaggcgtga gctaccatgc cccaccaaca ccccagtttt gtggaaaaga tgccgaaatt 2040

ccttttaag gagaagctga gcatgagcta tcttttgtct catttagtgc tcagcaggaa 2100

aatttgtatc tagtcccata agaacagaga gaggaaccaa gggagtggaa gacgatggcg 2160

ccccaggcct tgctgatgcc atatgccgga gatgagacta tccattacca cccttcccag 2220

caggctccca cgctcccttt gagtcaccct tcccagctcc agagaaggca tcactgaggg 2280

aggcccagca ccatggtcct ggctgacaca tggttcagac ttggccgatt tatttaagaa 2340

attttattgc tcagaacttt ccctccctgg gcaatggcaa gagcttcaga gaccagtccc 2400

ttggagggga cctgttgaag ccttcttttt ttttttttt aagaaataat cttgctctgt 2460

tgcccaggct ggagtgcagt ggcacaatca tagctcactg taacctggct caagcgatcc 2520

tcctgagtag ctaggactat aggcatgtca ctgcacccag ctaatttttt ttttttttt 2580

ttttttttt ttgcgacata gtctcgctct gtcaccaggc tggagtgcag tggcacgatc 2640

ttggctcact gcaacctctg cctcccgggt tcaagcaatt ttcctgcctc agcctcctga 2700

gtagctggga ctacaggcgc gtgtcaccac gcccagctaa tttttgtatt tttagtggag 2760

acagggtttc accatgttgg ctaggatggt ctcaatctct tgacctggtg atccatccgc 2820

cttggcctcc caaagtgcta ggattacagg cgtgagtcaa cctcaccggg catttttttt 2880

ttgagacgaa gtcttgctct tgctgcccaa gctggaatgt ggtggcatga tctcggctca 2940

ctgcaacctc cacctcctag gttcaagcga ttctccacct tagcctcccc agcagctggg 3000

attacaggtg cccatcaaca cacccggcta atttttgtat ttttattaga gatggggttt 3060

tgccatgttg gccaggctgc tctcgaactc ctaacctcag gtgatccacc cccattggcc 3120

tcccaaaata ctgggattac aggcatgagc caccgtgccc agctgaattt ctaaattttt 3180

gatagagatc gggtctttct atgttgccca agctggtctt gaactcctag cctaaagcag 3240

tcttcccacc tcggcctccc agagtgtttg gaatacgtgc gtaagccacc acatctgccc 3300

tggagcctct tgttttagag acccttccca gcagctcctg gcatctaggt agtgcagtga 3360

catcatggag tgttcgggag gtggccagtg cctgaagccc acaccggacc ctcttctgcc 3420

ttgcaggttg cctgcggaca cgctgggcct ctgtcctgat gctgctgagc tccctggtgt 3480

ctctcgctgg ttctgtctac ctggcctgga tcctgttctt cgtgctctat gatttctgca 3540

ttgtttgtat caccacctat gctatcaacg tgagcctgat gtggctcagt ttccggaagg 3600

tccaagaacc ccagggcaag gctaagaggc actgagccct caacccaagc caggctgacc 3660

tcatctgctt tgctttggca tgtgagcctt gcctaagggg gcatatctgg gtccctagaa 3720

ggccctagat gtggggcttc tagattaccc cctcctcctg ccatacccgc acatgacaat 3780

ggaccaaatg tgccacacgc tcgctctttt ttacacccag tgcctctgac tctgtcccca 3840

tgggctggtc tccaaagctc tttccattgc ccagggaggg aaggttctga gcaataaagt 3900

ttcttagatc aatca 3915

<210> 9
<211> 806
<212> DNA
<213> Homo sapiens

<220>

<221> CDS
<222> (48)..(536)

<400> 9
ggcacgaggg ttttctccgc gggcgcctcg ggcggaacct ggagata atg ggc agc 56
Met Gly Ser
1

acc tgg ggg agc cct ggc tgg gtg cgg ctc gct ctt tgc ctg acg ggc 104
Thr Trp Gly Ser Pro Gly Trp Val Arg Leu Ala Leu Cys Leu Thr Gly
5 10 15

tta gtg ctc tcg ctc tac gcg ctg cac gtg aag gcg gcg cgc gcc cgg 152
Leu Val Leu Ser Leu Tyr Ala Leu His Val Lys Ala Ala Arg Ala Arg
20 25 30 35

gac cgg gat tac cgc gcg ctc tgc gac gtg ggc acc gcc atc agc tgt 200
Asp Arg Asp Tyr Arg Ala Leu Cys Asp Val Gly Thr Ala Ile Ser Cys
40 45 50

tcg cgc gtc ttc tcc tcc agg tgg ggc agg ggt ttc ggg ctg gtg gag 248
Ser Arg Val Phe Ser Ser Arg Trp Gly Arg Gly Phe Gly Leu Val Glu
55 60 65

cat gtg ctg gga cag gac agc atc ctc aat caa tcc aac agc ata ttc 296
His Val Leu Gly Gln Asp Ser Ile Leu Asn Gln Ser Asn Ser Ile Phe
70 75 80

ggt tgc atc ttc tac aca cta cag cta ttg tta ggt tgc ctg cgg aca 344
Gly Cys Ile Phe Tyr Thr Leu Gln Leu Leu Leu Gly Cys Leu Arg Thr
85 90 95

cgc tgg gcc tct gtc ctg atg ctg ctg agc tcc ctg gtg tct ctc gct 392
Arg Trp Ala Ser Val Leu Met Leu Leu Ser Ser Leu Val Ser Leu Ala
100 105 110 115

ggt tct gtc tac ctg gcc tgg atc ctg ttc ttc gtg ctc tat gat ttc 440
Gly Ser Val Tyr Leu Ala Trp Ile Leu Phe Phe Val Leu Tyr Asp Phe
120 125 130

tgc att gtt tgt atc acc acc tat gct atc aac gtg agc ctg atg tgg 488
Cys Ile Val Cys Ile Thr Thr Tyr Ala Ile Asn Val Ser Leu Met Trp
135 140 145

ctc agt ttc cgg aag gtc caa gaa ccc cag ggc aag gct aag agg cac 536
Leu Ser Phe Arg Lys Val Gln Glu Pro Gln Gly Lys Ala Lys Arg His
150 155 160

tgagccctca acccaagcca ggctgacctc atctgctttg ctttggcatg tgagccttgc 596

ctaaggggc atatctgggt ccctagaagg ccctagatgt ggggcttcta gattaccccc 656

tcctcctgcc atacccgcac atgacaatgg accaaatgtg ccacacgctc gctctttttt 716

acacccagtg cctctgactc tgtccccatg ggctggtctc caaagctctt tccattgccc 776

agggagggaa ggttctgagc aataaagttt 806

<210> 10
<211> 163

<212> PRT
<213> Homo sapiens

<400> 10

Met Gly Ser Thr Trp Gly Ser Pro Gly Trp Val Arg Leu Ala Leu Cys
1 5 10 15

Leu Thr Gly Leu Val Leu Ser Leu Tyr Ala Leu His Val Lys Ala Ala
20 25 30

Arg Ala Arg Asp Arg Asp Tyr Arg Ala Leu Cys Asp Val Gly Thr Ala
35 40 45

Ile Ser Cys Ser Arg Val Phe Ser Ser Arg Trp Gly Arg Gly Phe Gly
50 55 60

Leu Val Glu His Val Leu Gly Gln Asp Ser Ile Leu Asn Gln Ser Asn
65 70 75 80

Ser Ile Phe Gly Cys Ile Phe Tyr Thr Leu Gln Leu Leu Leu Gly Cys
85 90 95

Leu Arg Thr Arg Trp Ala Ser Val Leu Met Leu Leu Ser Ser Leu Val
100 105 110

Ser Leu Ala Gly Ser Val Tyr Leu Ala Trp Ile Leu Phe Phe Val Leu
115 120 125

Tyr Asp Phe Cys Ile Val Cys Ile Thr Thr Tyr Ala Ile Asn Val Ser
130 135 140

Leu Met Trp Leu Ser Phe Arg Lys Val Gln Glu Pro Gln Gly Lys Ala
145 150 155 160

Lys Arg His

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US04/031481

International filing date: 23 September 2004 (23.09.2004)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/505,527
Filing date: 23 September 2003 (23.09.2003)

Date of receipt at the International Bureau: 24 November 2004 (24.11.2004)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse